# EZTitles Subtitling Assistant

*Quick Start Guide*

# Introduction to the EZTitles Subtitling Assistant

**Subtitling Assistant** is the next gen smart feature we've implemented into EZTitles intended to automate the process of creating Closed Captions both Pop-on and Roll-up, SDH Subtitles or Master Translation Templates in the same language. The **Subtitling Assistant** can be used in four ways:

1. Recognize the spoken text from the audio and automatically generate captions/subtitles in the same language. Extremely useful when creating captions or subtitles from scratch.
2. Split loaded text to subtitles and automatically synchronize (timecode) them with the audio. Useful if you already have a Dialogue List, Spotting List or any kind of transcript. Just load the text into EZTitles and with a single command you will have subtitles/captions perfectly timed with the audio. The only restriction for now is that the audio and the script languages should be the same.
3. Use the timing of an existing file or script and add the spoken lines transcribed from the audio for them.
4. Use the text of an existing transcription to automatically synchronize it to the audio.

Designing the Subtitling Assistant, we thought long and hard what are the best speech recognition technologies available and which are the leading Speech-to-Text solutions on the market. And more importantly, which of them are most likely to greatly improve in the future. So, we've decided to evolve EZTitles with the best tech available on the market, offered by one of the most innovative companies out there – Microsoft and AssemblyAI.

We've worked hard to get the best out of **Microsoft's and AssemblyAI's speech recognition technologies** and put them to good use in EZTitles adding our own expertise into the mix and adjusting everything to the high standards we've set for our software. Subtitling Assistant uses the power of AI technology to create perfectly timed captions. The Assistant also takes advantage of Natural Language Processing algorithms to optimize subtitle and row breaking.

The AssemblyAI algorithms we've implemented works great for Speaker change detection in English (US, UK and Australia).

And even better, **the video won't leave the security of your own computer during the process**.

The Subtitling Assistant provides excellent results for English audio and most of the western languages. Of course, we're pretty certain that would significantly improve in the future.

Documentaries or other types of content where the speech or narrative is clear will be transcribed almost perfectly. But transcription from noisy scenes, characters speaking in heavy accent or mumbling voices are not nearly as good.

So, all in all it can't completely automate the subtitling and very much so – the actual translation process for all scenarios. But we see it as significant step forward in taking away most of the hassle and letting you focus more on the creative stuff instead. What we're pretty confident of is the Subtitling Assistant will speed up your subtitling significantly.

# Getting started with the Subtitling Assistant

To start working with the Subtitling Assistant you should load the video file, wait for the Shot Changes detection and Audio Graph building to complete (*the "Processing video file" progress bar at the bottom right corner of EZTitles should fill up to the end*). After that you can click on:
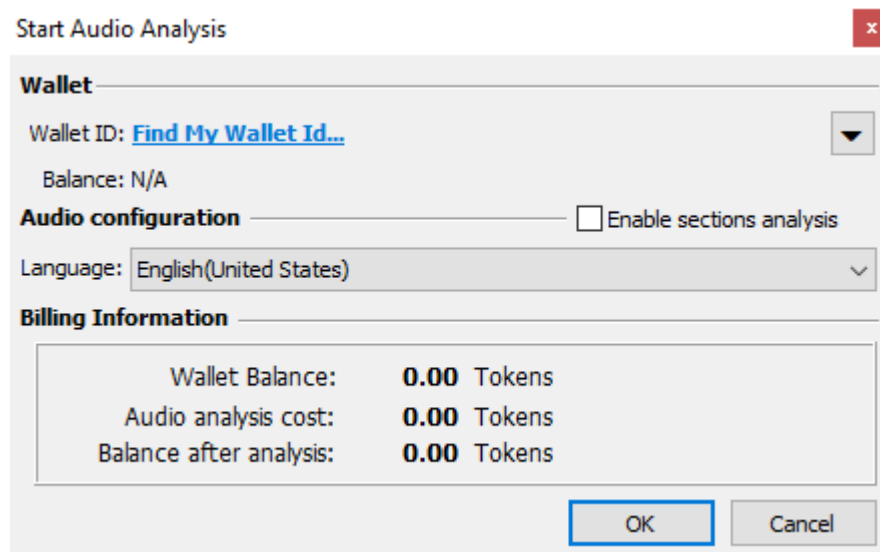**Subtitling Assistant->Start Audio Analysis**.

At this point you may see the following popping up on your screen:
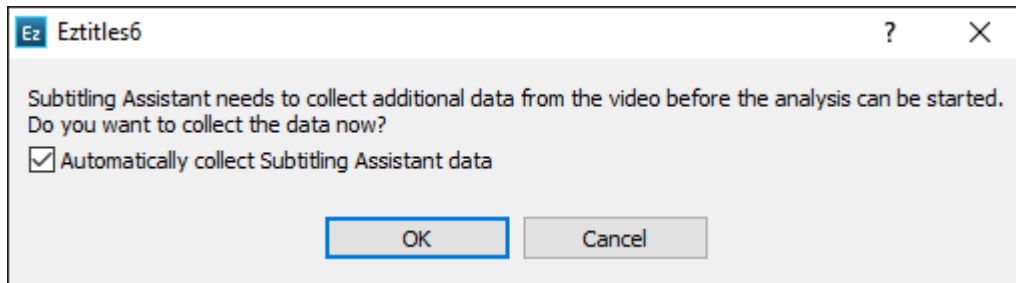


It appears because the Subtitling Assistant needs to collect more information about the audio. By ticking off "Automatically collect Subtitling Assistant data" EZTitles will do it automatically, while detecting the shot changes and building the audio graph, when a new video is opened.

Please, keep in mind that you can always deactivate the option from **Edit->Preferences->Performance**.

If the above pop-up window doesn't pop-up at all or if you've pressed the OK button already, you'll see the following window:
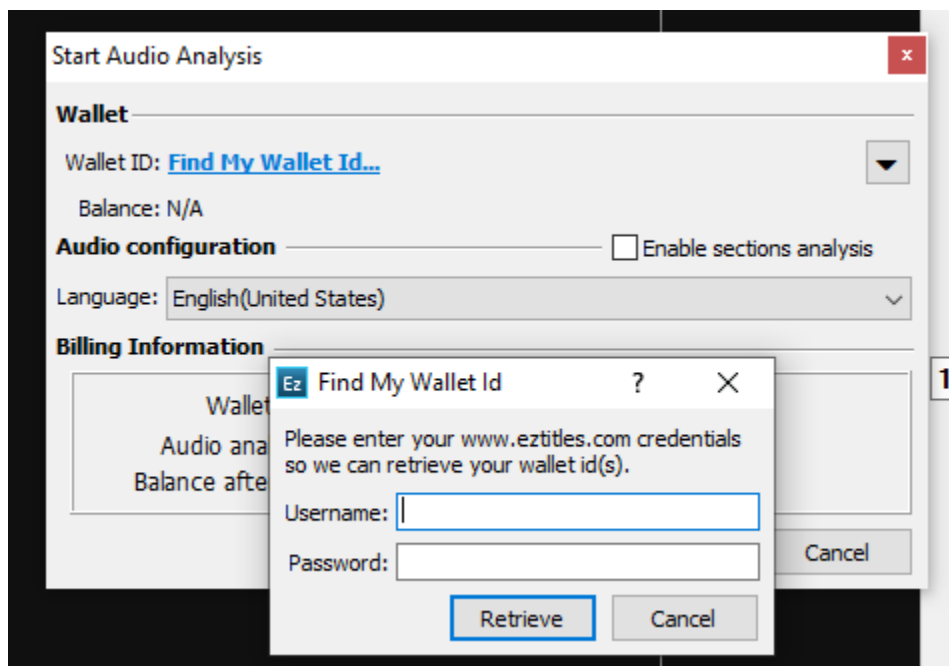
## Wallet

The Subtitling Assistant is a pay-as-you-go-based service and it requires **Tokens** to be loaded into a Wallet. The audio analysis costs certain amount of Tokens:
**1 Token is charged for each minute of audio analyzed.**

We also have a guide dedicated to managing the Tokens purchased and your Wallets.

The audio is analyzed just once for the selected language. All other operations that will be performed later on, like generating captions/subtitles or re-timing the already loaded in EZTitles third-party transcript, will not be charged.

To start working with Subtitling Assistant you have to enter a **Wallet ID**. If you are starting the Assistant for the first time click **Find My Wallet Id**. EZTitles will ask you to enter your **Customer username and password**.



After you enter your credentials EZTitles will show a list with all wallets currently available to your account. Choose one to use for your following audio analyses.

If you already have your Wallet ID, simply click on the arrow button next to **Wallet ID**, press on **Add Wallet** and then paste in the **Wallet ID** that we'll provide you with via e-mail:
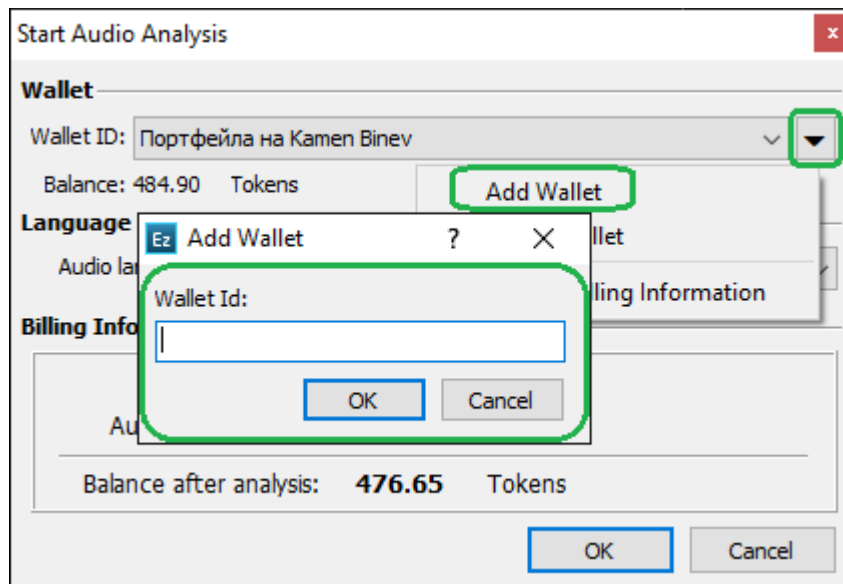


## Language

Right now, the Subtitling Assistant is working with same language transcriptions only – **the language of the original audio must be identical to the language of the subtitles/captions that will be created in the end.** Please, make sure to select the appropriate language from the **Audio Language** drop-down.

## Billing Information

The Billing Information section provides additional details about the Wallet in use and estimated costs for analyzing the currently loaded clip.

Now, after we've added a Wallet and configured the Language for the analysis, we can continue and start the process by pressing the **OK** button.

The audio starts uploading to the server. At this moment you could stop the process by using the **Stop Audio Analysis** command from the Subtitling Assistant menu. This may be necessary if you have made a mistake during the audio analysis configuration and have started it with wrong parameters. Once the audio upload has finished, you'll be charged for the audio analysis and the process could not be reversed.

# Creating Subtitles with the Assistant

The Subtitling Assistant complies with all configurable settings in EZTitles relevant to the timing of the subtitles. You can find more about these from the "Optimizing Subtitling Assistant Performance" topic in our User Guide.

Assuming the video has been processed already, the shot changes are detected and the audio is analyzed, you can continue with preparing the subtitles by using the "Subtitles from audio" option from the Subtitling Assistant drop-down menu:



Configuration options are split into 5 different sections and we'll take detailed look at each one of them.

# Assistant action

This section defines the results we want to achieve by using the Assistant and offers the choice between the tasks it can complete or scenarios it can be used in.

## Generate subtitles from audio

This is the most straight forward operation.

Using the audio analysis data acquired in the previous step EZTitles automatically generates subtitles and synchronizes them with the audio. The generated subtitles are in the same language as the audio of the video file.

The subtitles would need to be checked by the editor for inaccurately transcribed parts, for spelling and grammar mistakes and to pass the necessary quality control checks.

## Align loaded text to audio

Splits loaded text into subtitles and automatically synchronizes (timecodes) them with the audio.

Useful if you already have a Dialogue List, Spotting List or any kind of transcript. Just import the text into EZTitles and use this option. The only restriction for now is that the audio and the script should be the same language.

If there are differences between the loaded text and the transcribed audio (for example the audio hasn't been transcribed well enough), the Subtitling Assistant will try to compensate by automatically creating the best possible matching subtitles. These subtitles will be marked in green in the Preview list in order to be easily assessed by the editor.

## Extract transcription for subtitles

Consider this option if you already have well-synchronized empty subtitles and need to fill their text automatically.

Subtitling Assistant will not modify their timing but will replace their content with the spoken text recognized from the audio.

## Generate timecode for subtitles

This option can be used if you already have well-arranged subtitles without timecodes and you need just to synchronize them automatically.

Subtitling Assistant will not change the text in this case and will only set or modify the timing.

# Assistant Scope

All of the scope options define the range of operation of the Subtitling Assistant.

## Whole File

Subtitling Assistant will process the whole file. In case, the audio analysis is still going on, only the analyzed part(s) will be processed by the Assistant.

This operation requires some time to complete, sometimes even a couple of minutes if the video is particularly long.

### From current position to video end

Switch to Edit mode first, then select the subtitle you want to start with, position the video somewhere before this subtitle and then use this option.

If the align algorithm finds significant differences between the loaded text and the audio the process is stopped. Then, you should manually create subtitles for the mismatched part and then, when the script and the video are in sync again use this option.

### From current position to speech segment end

The speech segment is best described as a small part of the video with speech containing at least 50 words ending up with at least 2 sec pause in the speech.

Similar to the process described in the previous topic first select the subtitle you want to start with, position the video just before this subtitle and then use this option.

It works for **small parts of the video** and almost **instantly generates subtitles** for them. Then you can quickly review the subtitles, navigate to position in the video after the last subtitle and run the Assistant for next speech segment.

It can be really handy when adjusting and modifying the different **Subtitles Preferences** or when the audio isn't fully analyzed yet.

## Cues preferences

For now, there is just a single option **End of speech - outcue gap** that specifies the number of frames automatically added to the Out-cue of the last subtitle of a dialogue.
Typically, subtitles composing the dialogue are chained on the minimum interval and the last subtitle in the chain ends up several frames after the audio.

The default value is 12 frames but you can always **switch it off by specifying 0 frames**.

## Closed Captions

Specifies the style (Pop-on, Roll-up or Paint-on) for the automatically generated captions.

The option is only available when the project is configured in Closed Captions mode from the Project Settings menu.

## Subtitles preferences

The options in this section define the rules used to split the text into individual subtitles. All options except "Non-speech information in scripts" and their use are described in "Subtitling Assistant/Subtitles from Audio" topic of the Help (User's Guide) integrated in EZTitles.

Details about the "Non-speech information in scripts" option is provided below.

### Configurable rules with Weights

The advanced options, found by pressing the `>>` button can be used to increase the result's precision by modifying their weight to increase or decrease their priority respectively. More information about these rules could be found in the same topic of the Help (User's Guide).

# Non-speech information in scripts

The Non-speech Information options are available only when "Align loaded text" is enabled. The script's text can be searched for character names, sound effects, on-screen text and any other additional descriptive information not included in the spoken dialogue. Then you can decide if they should remain in the final version of the subtitles.

For now, we can distinguish between two types of "non-speech information":

**Speaker ID**: As the name suggests it is used to identify the speaker.

When the text is processed new speakers will be moved to new subtitle or to the beginning of new line.

**Other non-speech information (NSI):** includes sound effects and any other information not included in the speech.

NSI in the middle of the line will not be changed, as long as "Leave in subtitles" is clicked on:

"I (laughs) fully intend to!"

When the NSI is placed on a separate text line:

Good bye!
[Door Slams]

it will be moved to a new subtitle with zero duration because the Assistant cannot guess how long the event might be. We recommend doing Check against "Inconsistent Cues" to spot and identify all such subtitles.

## Non-speech information configuration

Non-speech Information configuration looks like this:
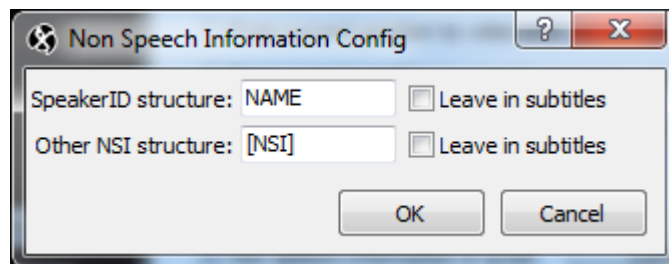


**Speaker ID** can be distinguished using the **NAME** or **name** wildcards which are case sensitive.

**NAME** will match any text in uppercase containing words longer than one symbol. The wildcard can be used alone or surrounded by other tags/symbols like: NAME:, [NAME], (NAME), etc.

For example, if we have the following:

**MAN IN CAR:** I've got something for you.
**JACK:** Oh, boy…

The Speaker ID can be configured like this:

**name** - the lowercase variant, will match any uppercase or lowercase letter. This variant must be surrounded by additional identification tags/symbols like parenthesis: (name), square brackets: [name] or anything else suitable in the particular case. Otherwise, the parameter will be ignored because it will match pretty much every word in the script.

Consider the following example where the Speakers are in sentence case:

> **[Man in Car]:** I've got something for you.
> **[Jack]:** Oh, boy...

the lowercase **name** wildcard shall be used instead:



**Other non-speech information** events or simply NSIs can be handled by the **NSI** case-insensitive wildcard.

The only limitation is that it must be surrounded by other identification tags/symbols to work properly like, for example, parenthesis: (NSI), square brackets: [NSI], hash-tag: #NSI#, etc.



Otherwise, if used alone, the parameter will be ignored because every word in the script will be matched.

### Leave in subtitles

Additionally, you can decide if the Speaker ID and/or other NSI will be kept in the text by selecting the "Leave in subtitles" option.

# Intelligent Commands

To complement the Subtitling Assistant to further improve the level of automation in EZTitles we have also developed something that no other application offers at the moment - **Intelligent Commands**.

**Split**, **Merge** and **Move Text to Previous/Next subtitle** commands now benefit from the Audio Analysis of the video and will automatically adjust the subtitle's timing according to the exact moment the line, word or phrase is spoken in the movie.